

METHODS FOR DETECTION OF NUCLEIC ACID
POLYMORPHISMS USING PEPTIDE-LABELED
OLIGONUCLEOTIDES AND ANTIBODY ARRAYS.

1. FIELD OF THE INVENTION

5 The present invention relates to compositions and methods for use in screening nucleic acid populations for nucleotide polymorphisms. The methods, referred to generally as ValiGeneSM Mutation Screening, Peptide-Linked (VGMS-PL) methods, are specifically designed for high-throughput genotype mapping and gene expression analysis of nucleic acids without requiring a PCR amplification step. In particular, the methods of
10 the invention utilize oligonucleotide probes labeled with distinguishable and identifiable labels, (e.g., peptide tags), that are captured on addressable antibody arrays for analysis (e.g., by fluorescence photometry).

2. BACKGROUND OF THE INVENTION

15 With the advent of genome-wide sequencing efforts, understanding the molecular basis of all genetic diseases may soon be within reach. Single nucleotide polymorphism (SNP) detection analysis is playing an increasingly powerful role in mapping out the underlying genetic basis of many human diseases.

 Approximately 1 in every 1000 nucleotides differs between any two copies
20 of the human genome (Cooper, 1996, Hum. Genet. 69:201-205). Some of these genetic variations, or SNPs, lead to differences in the proteins encoded by such genes. Others are "silent", residing in non-protein coding regions of the genome. Such SNPs are now being used, for example, to diagnose genetic disorders, determine a predisposition to genetic disease, identify or determine the ancestry of a genetic sample, or correlate genetic
25 sequences with phenotypic conditions, such as complex disorders or drug response and toxicity (Risch and Merikangas, 1996, Science 273:1516-1517). This powerful combination of genetic and molecular biological approaches is changing the face of drug development. SNPs have been correlated with Huntington's disease, Alzheimer's disease, and various forms of breast cancer. In the emerging field of pharmacogenomics, specific
30 SNPs are being used to determine and predict a patient's susceptibility to diseases as well as drug toxicity and response. Pharmacogenomics can also provide tools to identify new targets for designing drugs and to optimize the use of existing drugs. The hope is that this understanding will ultimately lead to the early diagnosis, prevention, and treatment of genetic diseases.

35 Single nucleotide polymorphisms can be identified by a number of methods, including DNA sequencing, restriction enzyme analysis, or site-specific hybridization.

However, high-throughput genome-wide screening for SNPs and other variations requires the ability to simultaneously analyze multiple loci with high accuracy and sensitivity. To increase sensitivity, current high-throughput methods for single nucleotide detection rely on a step that involves amplification of the target nucleic acid sample, usually by the polymerase chain reaction (PCR) (see, e.g., Nikiforov *et al.*, U.S. Patent No. 5,679,524 issued Oct 21, 1997; McIntosh *et al.*, PCT publication WO 98/59066 dated December 30, 1998; Goelet *et al.*, PCT publication WO 95/12607 dated May 11, 1995; Wang *et al.*, 1998, Science 280:1077-1082; Tyagi *et al.*, 1998, Nature Biotechnol. 16:49-53; Chen *et al.*, 1998, Genome Res. 8:549-556; Pastinen *et al.*, 1996, Clin. Chem. 42:1391-1397; Chen *et al.*, 1997, Proc. Natl. Acad. Sci. 94:10756-10761; Shuber *et al.*, 1997, Hum. Mol. Gen. 6:337-347; Liu *et al.*, 1997, Genome Res. 7:389-398; Livak *et al.*, Nature Genet. 9:341-342; Day and Humphries, 1994, Anal. Biochem. 222:389-395). However, the fidelity of the PCR technique is limited. Combinations of pairs of PCR primers tend to generate spurious reaction products. Moreover, the number of errors in the final reaction product increases exponentially with the each round of PCR amplification after an error is introduced into a DNA sample. Thus, PCR error can be a substantial drawback when searching for rare variations in nucleic acid populations.

For all of the reasons addressed above, a highly sensitive, highly specific, PCR-free method for high-throughput detection of nucleic acid variations is urgently needed. This invention provides such a method, as described in detail below.

Citation or discussion of a reference herein shall not be construed as an admission that such reference is prior art to the present invention.

3. SUMMARY OF THE INVENTION

The present invention provides methods for detection of single nucleotide polymorphisms (SNPs) and other variations in nucleic acid populations. The methods of the present invention for high-throughput PCR-free screening are used for detection of alterations and polymorphisms as well as for analysis of gene expression, both qualitative and quantitative, directly from cellular total RNA. The methods may be used, for example, to diagnose disorders, determine predisposition to genetic diseases, determine identity or ancestry, or correlate genetic sequences with phenotypic conditions.

In a specific embodiment, the invention relates to methods for efficient, sensitive, high-throughput addressable-array based screens using SNP-specific oligonucleotide probes having distinguishable and identifiable peptide tags to detect polymorphisms in nucleic acid target molecules. First, SNP-specific peptide-linked oligonucleotide probes, comprising distinguishable markers, are hybridized to a target nucleic acid sample. Next, any hybrid molecules formed are captured on high-density addressable antibody arrays and processed by enzymes that recognize and cleave the

captured hybrid molecules at mismatched base pairs. Finally, markers present on the cleaved hybrid molecules are then detected and analyzed to identify any polymorphic site(s) within a specific target nucleic acid molecule of interest.

The present methods offer several advantages over the currently available technologies for genotype detection. First, the methods described herein allow detection of genetic variation using minimal amounts of genetic material without requiring a PCR
5 amplification step, avoiding the introduction of new mutations into the sample being tested. In other genotyping methods, a PCR amplification step is typically used to amplify the signal of a given target sequence within a nucleic acid sample to allow detection. In the present invention, it is the signals that are amplified from a number of limited targets. This
10 allows reliable SNP detection using minimal amounts of genetic material from a variety of biological sources, such as biopsies of tissue from a patient with a potential genetic disorder. Second, unlike in other methods of genotype mapping, nucleic acid hybridization takes place in solution, eliminating the need to immobilize a hybridization partner. Solution hybridization is more efficient than hybridization with one immobilized partner,
15 resulting in increased efficiency of mismatch detection. Third, the addressable chip array allows a flexible detection system for high-throughput genotype analysis. Multiple SNP sites can be screened simultaneously from a patient or genetic sample, or alternatively, a single SNP site in many different DNA samples, can be tested simultaneously. For example, in one embodiment, detection of multiple polymorphisms in a target nucleic acid
20 sample is possible in a single run. Multiple probes can be individually prepared, each probe having a unique peptide label and a sequence corresponding to a polymorphic site to be detected. Such multiple probes can be hybridized in "batch" with the target nucleic acid sample. In another embodiment, multiple target nucleic acid samples can be screened simultaneously for the presence or absence of a single SNP locus.

Throughout this application reference is made to peptide labels and
25 antibodies for binding said labels. In addition to peptide-antibody combinations, it will be understood by those skilled in the art that any label can be used, in combination with a suitable binding partner. Examples of such labels and binding partners include, but are not limited to, digoxigenin-antidigoxigenin, biotin-streptavidin, ligand (*e.g.* hormone)-receptor,
30 and carbohydrate-lectin combinations. The invention may also be practiced using a nucleic acid label and a nucleic acid binding partner, wherein the label and the binding partner recognize one another by hybridization of a complementary sequence. Nucleic acid derivatives may also be used as labels and binding partners (*e.g.* peptide nucleic acids).

The term "polymorphism" as used herein refers to the presence of two or
35 more alternative genomic nucleotide sequences or alleles between or among different genomes or individuals. "Polymorphic" refers to the condition in which two or more variants of a specific genomic sequence can be found in a population. A "polymorphic site"

is the locus at which the variation occurs. A single nucleotide polymorphism, or SNP, is a single base pair variant, typically the substitution of one nucleotide by another nucleotide at the polymorphic site. Deletion or insertion of a single nucleotide also gives rise to a SNP.

While the terms "SNP" and "SNP detection" are used throughout this disclosure for purposes of clarity and simplicity of description, it is to be understood that these terms are meant to encompass methods for detection of double and multiple nucleotide polymorphisms as well as single nucleotide polymorphisms. Further, the terms "SNP" and "SNP detection" are meant to encompass methods for detection of deletions and insertions in nucleic acid sequences. Accordingly, substitutions, deletions or insertions of various sizes may be detected using the methods of the invention.

The SNP detection methods described herein utilize nucleotide polymorphism-specific probes. The nucleotide polymorphism-specific probe comprises one or more distinguishable "markers". A marker is a nucleotide residue that is incorporated into the probe, preferably during synthesis of the oligonucleotide, that either is 1) covalently bound to a detectable label (e.g. a fluorophore), or 2) is covalently bound to an affinity group (e.g. biotin) that is labeled post-synthesis of the nucleic acid by contacting said affinity group with a labeled cognate binding partner. A detectable label may include but is not limited to a luminescent compound, a chromophore, a fluorescent compound, a radioactive isotope or group containing same, or a nonisotopic label, such as an enzyme or dye. Thus, a detectable label may be directly linked to a nucleotide or indirectly linked, e.g., by its presence on a partner molecule that binds to an affinity group directly linked to the nucleotide.

The invention provides a method for high-throughput nucleotide polymorphism analysis of a nucleic acid sample from a subject comprising contacting a plurality of peptide-labeled oligonucleotide probes with the nucleic acid sample in solution, under conditions conducive to hybridization of the probes to nucleic acid in the sample; and detecting one or more probes of the plurality that hybridize to nucleic acid in the sample using an antibody array comprising antibodies immunospecific to one or more of the peptide labels. In one embodiment of this method, the detecting one or more probes is carried out without using PCR or MutS, an *E. coli* mismatch binding protein that recognizes and binds to nucleic acids containing mismatched base pairs.

The invention further provides a method for screening a nucleic acid sample from one or more subjects for the presence of a polymorphism comprising the following steps in the order stated: a) contacting the nucleic acid sample in solution with one or more nucleotide polymorphism-specific peptide-labeled oligonucleotide probes, under conditions conducive to hybridization of the probes to nucleic acid in the sample; each probe comprising a first marker covalently attached to a first detectable label, and a second marker covalently attached to a second detectable label that produces a signal distinguishable from

the first detectable label; such that one or more hybrid molecules are formed between the nucleic acid and one or more oligonucleotide probes; b) capturing at least one of the one or more hybrid molecules on a solid phase surface; c) contacting the solid phase surface with mung bean nuclease, under conditions wherein the nuclease is active; d) removing material not bound to the solid phase surface; e) detecting or measuring from the solid phase surface a first signal from the first detectable label and a second signal from the second detectable label; f) cleaving the hybrid molecules on the solid phase surface at mismatched base pairs; and g) detecting or measuring from the solid phase surface a third signal from the first detectable label and a fourth signal from the second detectable label, determining a first ratio of the second signal to the first signal, and a second ratio of the fourth signal to the third signal, and comparing the first ratio to the second ratio, wherein a difference between the first ratio and the second ratio indicates that a polymorphism is identified.

In another embodiment, the invention provides a method for screening a nucleic acid sample from one or more subjects for the presence of a polymorphism comprising the following steps in the order stated: a) contacting the nucleic acid sample in solution with one or more nucleotide polymorphism-specific peptide-labeled oligonucleotide probes under conditions conducive to hybridization of the probes to nucleic acid in the sample; each probe comprising a first and a second marker; such that one or more hybrid molecules are formed between the nucleic acid and one or more nucleotide polymorphism-specific oligonucleotide probes; b) capturing at least one of the one or more hybrid molecules on a solid phase surface; c) contacting the solid phase surface with mung bean nuclease, under conditions wherein the nuclease is active; d) removing material not bound to the solid phase surface; e) contacting the solid phase surface with (i) a first partner molecule with the ability to specifically bind the first marker, and (ii) a second partner molecule with the ability to specifically bind the second marker, said first partner molecule comprising a first detectable label and said second partner molecule comprising a second detectable label that produces a signal distinguishable from the first detectable label; f) removing material not bound to the solid phase surface; g) detecting or measuring from the solid phase surface a first signal from the first detectable label and a second signal from the second detectable label; h) cleaving the hybrid molecules on the solid phase surface at mismatched base pairs; and i) detecting or measuring from the solid phase surface a third signal from the detectable label and a fourth signal from the second detectable label, determining a first ratio of the second signal to the first signal, and a second ratio of the fourth signal to the third signal, and comparing the first ratio to the second ratio, wherein a difference between the first ratio and the second ratio indicates that a polymorphism is identified. In one embodiment of these methods the first or second marker is covalently attached to a biotin moiety and the first or second partner molecule is avidin or streptavidin. In another embodiment, the method of claim 4 wherein the first or second marker is

covalently attached to a carbohydrate moiety and the first or second partner molecule is a lectin. In another embodiment, the first partner molecule is an antibody that binds specifically to the first marker and the second partner molecule is an antibody that binds specifically to the second marker.

In another embodiment, a method is provided for screening a nucleic acid sample from one or more subjects for the presence of a polymorphism comprising the following steps in the order stated: a) contacting the nucleic acid sample in solution with one or more nucleotide polymorphism-specific peptide-labeled oligonucleotide probes, under conditions conducive to hybridization of the probes to nucleic acid in the sample; each probe comprising a first and a second marker; such that one or more hybrid molecules are formed between the nucleic acid and one or more nucleotide polymorphism-specific oligonucleotide probes; b) capturing at least one of the one or more hybrid molecules on a solid phase surface; c) contacting the solid phase surface with mung bean nuclease, under conditions wherein the nuclease is active; d) removing material not bound to the solid phase surface; e) contacting the solid phase surface with a first primary partner molecule, with the ability to bind the first marker, and a second primary partner molecule with the ability to bind the second marker; f) removing material not bound to the solid phase surface; g) contacting the solid phase surface with (i) a first secondary partner molecule, with the ability to bind the first primary partner molecule, said first secondary partner molecule comprising a first detectable label, and (ii) a second secondary partner molecule, with the ability to bind the second primary partner molecule, said second secondary partner molecule comprising a second detectable label that produces a signal distinguishable from the first detectable label; h) removing material not bound to the solid phase surface; i) detecting or measuring from the solid phase surface a first signal from the first detectable label and a second signal from the second detectable label; j) cleaving the hybrid molecules on the solid phase surface at mismatched base pairs; and k) detecting or measuring from the solid phase surface a third signal from the first detectable label and a fourth signal from the second detectable label, determining a first ratio of the second signal to the first signal, and a second ratio of the fourth signal to the third signal, and comparing the first ratio to the second ratio, wherein a difference between the first ratio and the second ratio indicates that a polymorphism is identified. In one embodiment, the first or second marker is covalently attached to a biotin moiety and the first or second primary partner molecule is avidin or streptavidin. In another embodiment, the first or second marker is covalently linked to a carbohydrate moiety and the first or second primary partner molecule is a lectin. In yet another embodiment, the first and second primary partner molecules are distinct primary antibodies, and the first and second secondary partner molecules are distinct secondary antibodies.

The invention further provides a composition comprising one or more nucleotide polymorphism-specific peptide-labeled oligonucleotide probes bound to one or more antibodies of an antibody array.

In one embodiment of these methods, at least one of the one or more nucleotide polymorphism-specific peptide-labeled oligonucleotide probes comprises: a) a peptide covalently attached to the 5' end of the oligonucleotide; b) a first marker at the penultimate 5' position of the oligonucleotide; and c) a second marker at the 3'-end of the oligonucleotide.

In another embodiment of these methods, the solid phase surface comprises a plurality of loci, wherein each locus is capable of specifically binding to one of the one or more oligonucleotide probes via the peptide of the peptide-labeled oligonucleotide and wherein the peptide is covalently attached to the 5' end of the oligonucleotide. The method of claim 4 wherein the first or second marker is covalently attached to a carbohydrate moiety and the first or second partner molecule is a lectin.

In another embodiment, either or both the first detectable label or the second detectable label is an enzyme, a fluorophore, a chemiluminescent label, or a radioisotope. In another embodiment, either or both the first detectable marker or the second detectable marker is a fluorophore.

In various embodiments of the invention, the signals are measured by comparing the first ratio to the second ratio, wherein the first ratio at least 35% greater than the second ratio indicates that a polymorphism is identified.

In one embodiment, the solid phase surface comprises a plurality of loci, wherein each locus comprises an antibody specific to one or more of the peptides of the peptide-labeled oligonucleotide probes. In another embodiment, the solid phase surface is a plastic chip. In another embodiment, the solid phase surface is the well of a microtiter plate. In another embodiment, the nucleic acid in the sample is less than 2 μ g. In another embodiment, the nucleic acid in the sample is selected from the group consisting of genomic DNA and cDNA. In another embodiment, the nucleic acid in the sample is genomic DNA. In another embodiment, the nucleic acid in the sample is cDNA. In another embodiment, the nucleic acid in the sample is double stranded DNA. In another embodiment, the nucleic acid in the sample is single stranded DNA. In another embodiment, cleaving the hybrid molecules at mismatched base pairs is carried out by contacting said hybrid molecules with one or more specific nucleases under conditions that allow cleavage of said hybrid molecules at mismatched base pairs. In another embodiment, cleaving the hybrid molecules at mismatched base pairs is carried out by contacting said hybrid molecules with *E. coli* endonuclease V and S1 nuclease under conditions that allow cleavage of said hybrid molecules at mismatched base pairs. In another embodiment, the

first and second detectable labels are fluorophores that absorb at the same wavelength but emit at a distinct frequency.

In various embodiments, the invention encompasses methods for detecting or measuring the presence of an alternatively spliced RNA transcript in a that an alternatively spliced RNA transcript is in the sample.

5 In one embodiment, the subject is a plant. In another embodiment, the subject is a virus, a bacterium, a yeast, or a fungus. In another embodiment, the subject is a mammal. In another embodiment, the subject is equine, porcine, ovine, bovine, canine, feline, or human. In another embodiment, the subject is a human. In another embodiment, the subject is a plurality of human subjects that exhibit a phenotype of interest.

10 It is to be understood, that while a SNP-specific probe is represented throughout this application as having two markers, SNP probes comprising three, four, or more differentially-labeled markers are also within the scope of the invention. For example, such markers may be placed at locations within the sequence of an oligonucleotide to allow detection of alternate mismatched base pairs or splice junction sequences.

15 The methods of the invention can be used for direct, PCR-free, genotype mapping, as well as quantitative genomic analysis of allele amplification and loss-of heterozygosity phenomena. The invention further provides a method for detecting and measuring RNA in a sample from a subject comprising the method described above, wherein the SNP-specific oligonucleotide probe comprises the sequence of a known RNA
20 transcript or a known splice-site junction sequence. Such methods can be used to monitor gene expression, both qualitatively and quantitatively, in very small amounts of RNA sample. In one embodiment, the sample is total cellular RNA extracted from a tissue sample or biopsy. In another embodiment, such methods are used *in situ* to monitor gene expression within histological preparations for diagnostic and prognostic purposes.

25 4. BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1. Schematic representation of the VGMS-PL method.

FIG. 2. Method for genotype mapping. This method may use positive and negative oligonucleotide probes, each labeled with a peptide (open box) and two distinct
30 fluorophores (F_1 and F_2).

FIG. 3. Schematic representation of the method for gene expression analysis.

FIG. 4. A two-part peptide-labeled oligonucleotide probe. The probe illustrated is labeled with a peptide (open box) and two distinct fluorophores (F_1 and F_2) on
35 either side of a polymorphism (N) to be detected. The location of nucleotide base derivatives resistant to nuclease cleavage (e.g. thiolated bases) is indicated by shaded rectangles.

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention provides methods referred to generally as ValiGeneSM Mutation Screening, Peptide-Linked methods (VGMS-PL methods) for identifying polymorphisms present in populations of nucleic acids. The methods may be used for, for example, determining drug response and toxicity, diagnosing genetic disorders, determining predisposition and risk factors for genetic diseases, determining an individual's identity or ancestry, and correlating genetic sequences with physical disorders or syndromes.

In currently available genotyping methods, a PCR step is typically used to increase the amount of a given target within a genomic sample by amplification of the entire sample to increase its detection. In the present invention, detection is increased by amplification of the signals from a number of limited targets, obviating the need for a PCR amplification step. Signal amplification is achieved by incorporation of affinity groups into the peptide-labeled oligonucleotide probe, and, in later steps, contacting said affinity groups with partner molecules having detectable labels.

The present invention utilizes solution hybridization and cleavage by endonucleases to recognize and detect mismatches. The invention thereby permits avoidance of the use of *E. coli* mismatch binding proteins such as MutS, an *E. coli* mismatch binding protein that recognizes and binds to nucleic acids containing mismatched base pairs, which has been used in the past for mismatch detection methods (Wagner et al., 1995, Nucl. Acids Res. 23:3944-3948; Wagner et al. EPO 596 028 B1 and PCT 94/08022).

Specifically, the methods disclosed herein below employ nucleotide polymorphism-specific probes, comprising peptide-linked oligonucleotides having distinguishable markers, as hybridization probes for target nucleic acid molecules. The basic working principle of the PCR-free direct SNP detection method is outlined in Figure 1. A nucleic acid sample, taken from an individual subject or a population of individual subjects, is annealed to a peptide-linked oligonucleotide probe specific for a genetic locus of interest. For example, such a locus can be a SNP, where the nucleic acid sample is genomic DNA. In an alternative embodiment, where the nucleic acid sample being assayed is RNA, the locus can be a splice site. The annealed probe:target nucleic acid hybrid is then trapped on a solid phase surface, and enzymatically treated to eliminate free probe. A preferred, but optional, signal amplification step is performed, prior to the reading the signals from the probe's 5'- and 3'- end distinguishable markers. The solid phase surface is then treated with a set of enzymes that specifically digests the hybrids at mismatched basepairs, and a second signal is recorded. Analysis of the readouts from the hybrids before and after their digestion at mismatched base pairs allows one to determine the presence of, and in various embodiments, precisely quantitate, the presence of a SNP or splice site in the genetic loci of interest. In one embodiment, the hybrid molecules are captured on high density addressable arrays and analyzed by enzymes that recognize and cleave the captured hybrid molecules at

mismatched base pairs. The detectable labels present on the cleaved hybrid molecules are then detected and analyzed to identify polymorphic sites within the specific target nucleic acid molecule of interest.

In the various embodiments of the present invention, labeled oligonucleotides are used as polymorphism-specific hybridization probes. In addition to a peptide label used for capturing a probe on an addressable antibody array, such probes are also labeled with markers to detect hybridization with target nucleic acids at various steps of the process. As noted above, the basic working principle of the VGMS-PL methods is set forth in Figure 1. A schematic diagram of PCR-free genotype mapping analysis is set forth in Figure 2. Representative positive and negative probes are illustrated in the top panel of Figure 2. Details of the method are described in detail herein below.

5.1 OLIGONUCLEOTIDE PROBES

In various embodiments of the present invention, peptide-labeled SNP-specific or locus-specific oligonucleotides are used as hybridization probes to target and detect polymorphic sequences. In a preferred embodiment, an SNP-specific probe comprises a synthetic oligonucleotide with three properties (a positive and negative probe is depicted in the top panel of Figure 2). First, a peptide label is linked to the 5' end of the oligonucleotide (indicated by the open box in Figure 2). Second, a first marker (F_1), which may be a detectable label or, alternatively, an affinity label that can bind to a partner molecule such as an antibody or affinity partner linked to such a first marker, is located on the penultimate deoxyribose base at the 5' end of the oligonucleotide (indicated by F_1 in Figure 2). Third, a second marker (F_2), distinguishable from the first marker, is located on the last deoxyribose base at 3' end of the oligonucleotide (indicated by F_2 in Figure 2).

Polymorphism-specific oligonucleotide probes may be designed to correspond to any nucleic acid sequence, preferably a region known or suspected to contain a polymorphism. Numerous clinically important polymorphisms are known. For example, SNPs in apolipoprotein genes, such as apolipoprotein E variants, are implicated in diseases such as atherosclerosis (see Breslow, 1996, *Science* 272:685-688) and Alzheimer's disease (Selkoe, 1997, *Science* 275: 630-631; Weisgraber, 1996, *FASEB J.* 10:1485-1494).

Mutation in apolipoproteins and the low density lipoprotein (LDL) receptor gene is common in patients with familial hypercholesterolemia (Hobbs *et al.*, 1990, *Ann. Rev. Genet.* 24:133-170; Brown *et al.*, 1984, *Sci. Amer.* 251:58-66). Polymorphisms in the β -globin gene locus are also well known; for example, a single base pair mutation in the β -globin gene is responsible for the substitution of the amino acid valine in place of glutamic acid, resulting in sickle cell anemia (Ingham, 1957, *Nature* 180:326-328).

Any gene known or suspected to contain a polymorphic site may be used as a target nucleic acid in the methods of the present invention. Amino acid sequences and

nucleotide sequences of known or suspected SNP sites are generally available in sequence databases, such as GenBank, or the HUGO database initiative, which is establishing a mutational database as a source of common variants of human disease (Cotton *et al.*, 1998, Science 279:10-11). Computer programs, such as Entrez, can be used to browse the database and retrieve any nucleotide or amino acid sequence data of interest (see
5 <http://www.ncbi.nlm.nih.gov/Entrez>). These databases can also be searched to identify sequences with various degrees of similarities to a query sequence using programs, such as FASTA and BLAST, which rank the similar sequences by alignment scores and statistics.

The SNP-specific probe comprises one or more distinguishable "markers". A marker is a nucleotide residue that is incorporated into the probe, preferably during
10 synthesis of the oligonucleotide (see Section 5.4.3), that either is 1) covalently bound to a detectable label (*e.g.* a fluorophore) or, 2) is covalently bound to an affinity group (*e.g.* biotin) that is labeled post-synthesis of the nucleic acid by contacting said affinity group with a labeled cognate binding partner. A detectable label may include but is not limited to
15 a luminescent compound, a chromophore, a fluorescent compound, a radioactive isotope or group containing same, or a nonisotopic label, such as an enzyme or dye. Thus, a detectable label may be directly linked to a nucleotide or indirectly linked, *e.g.*, by its presence on a partner molecule that binds to an affinity group directly linked to the nucleotide. For example, affinity groups or partner molecules that may be used include, but are not limited to biotin, avidin, streptavidin, digoxigenin, haptens, and monoclonal and polyclonal
20 primary or secondary antibodies.

Where two or more markers are present on a single SNP-specific probe, it is important that the two markers are distinguishable. Assays for such distinguishable markers are well known in the art (see, *e.g.*, Harlow and Lane, 1988, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, Chapter
25 14). In a preferred embodiment, this is accomplished by the use of distinguishable partner molecules that bind to different affinity groups that are linked to the marker residue of the probe. For example, the 3' marker of the SNP-specific probe may be linked to a biotin molecule. Streptavidin, a binding partner molecule of biotin, may be linked to an enzyme label, such as horse radish peroxidase (HRP), and used as an affinity binding partner to
30 specifically bind to the biotin-linked marker. To differentially label the markers, the 5' marker of the SNP-specific probe may be linked to a chelate of divalent or trivalent cations (Ni^{2+} , Cu^{2+} , Eu^{3+} , etc., Dahlen *et al.*, 1994, Bioconjug. Chem. 5:268-72; Li and Selvin, 1997, Bioconjug. Chem. 8:127-132), and a PDLs-coupled globular proteins such as ovalbumin, thyroglobulin *etc.*, and used to specifically bind the chelate marker. PLDS, a polymer of
35 lysine, dipicolinate, and succinate, binds divalent and trivalent cations with high affinity and strongly enhances the visible luminescence they emit when excited with ultraviolet light (Lamture and Wensel, 1995, Bioconjug. Chem. 6:88-92).

In another embodiment, distinguishable markers of a specific probe comprise different fluorophores that emit at different frequencies. Details of the methods used for labeling, signal amplification, and signal detection are described in Section 5.4.

In the most preferred embodiment, SNP-specific probes comprise two distinguishable, detectable affinity labels. In another embodiment, the SNP-specific probe comprises three, four, or more differentially labeled markers. Such markers may be placed at locations within the sequence of an oligonucleotide to allow detection of alternate mismatched base pairs or splice junction sequences.

In various embodiments of the invention, more than one probe may be used in a single reaction to target more than one polymorphism along the same molecule. Where a mixture of probes is to be used in a single reaction, prior analysis is recommended to minimize interactions among probes and among target fragments, and to ensure similar ΔG characteristics among probes. Such prior analysis of oligonucleotide probes is well known in the art.

5.2 METHODS FOR PCR-FREE GENOTYPE MAPPING

In this embodiment, the methods of the invention are used to detect polymorphic sites, *e.g.*, SNPs, in a DNA sample. The methods may be used to detect single mismatches between a SNP-specific oligonucleotide probe and a target molecule, or to distinguish between single mismatches, double mismatches, or multiple mismatches.

Generally, the method for PCR-free genotype mapping of a DNA sample comprises the following steps: 1) contacting a target DNA with a SNP-specific peptide-labeled oligonucleotide probe comprising a first and a second marker; 2) capturing the annealed hybrids on a solid phase surface; 3) treating the solid phase surface with mung bean nuclease; 4) removing material not bound to the solid phase surface; 5) detecting a signal from the first and second markers; 6) cleaving the probe DNA at any mismatched base pair present; and 7) detecting another signal from the first and second markers.

In the first step, the target DNA is prepared from the sample. The target DNA can be obtained from a biological sample, such as, but not limited to, whole blood, plasma, serum, skin, saliva, urine, lymph fluid, cells obtained from biopsy aspirate, tissue culture cells, media, or non-biological samples such as food, water, or other material. Methods for preparation of DNA from such sources are well known to those of skill in the art (see, *e.g.*, Current Protocols in Molecular Biology series of laboratory technique manuals, 1987-1994 Current Protocols, 1994-1997 John Wiley and Sons, Inc.).

In one embodiment, the target DNA is a fragment of genomic DNA. In another embodiment, the test DNA is cDNA. In another embodiment, the target DNA is a mixed population of genomic DNAs derived from a plurality of subjects of interest, for example, subjects afflicted with a particular disorder.

Target DNA is prepared for hybridization by fragmenting the DNA to an average length of 300 base pairs. In a preferred embodiment, this is accomplished by treating with one or more restriction endonucleases prior to annealing with the probe, to generate fragments of appropriate length. The restriction enzymes are chosen to recognize and cleave a recognition site that does not appear within any SNP-specific probe being used. Such treatment preferably yields DNA fragments in the size range of approximately 100 to 1000 base pairs, more preferable in the range of 200 to 500 base pairs, most preferably, approximately 300 base pairs. Many restriction enzymes, preferably restriction enzymes with four-base recognition sites, are known in the art and commercially available and can be used for this purpose. In an alternative embodiment, DNA is fragmented by sonication to shear DNA molecules to an average size of 300 basepairs.

The DNA fragments are then mixed with peptide-labeled SNP-specific oligonucleotide probes under conditions that allow hybridization to occur. Exemplary conditions for hybridizations are as follows. Preferably, 300 ng of genomic DNA (about 9×10^4 copies of target gene containing SNP of interest) are mixed with 100 fg of probes (about 3.5×10^6 copies of probe) in a total volume of 5 μ l in a buffer containing 10 mM Tris-HCl (pH 7.5); 1 mM EDTA, 50 mM NaCl; and 1 mM cetyltrimethylammonium bromide (CTAB). The mixture is denatured, by heating to 98°C for 5 minutes, and allowed to renature by incubating the mix at 68 °C for 20 minutes. To optimize the results, the stringency can be adjusted. Such methods for adjustment of hybridization stringency are routine and well known in the art (*see, e.g.*, Sambrook *et al.*, 1989, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Ausubel *et al.*, eds., in the Current Protocols in Molecular Biology series of laboratory technique manuals, © 1987-1994 Current Protocols, © 1994-1997 John Wiley and Sons, Inc; Dyson, N.J., 1991, Immobilization of nucleic acids and hybridization analysis, In: Essential Molecular Biology: A Practical Approach, Vol. 2, T.A. Brown, ed., pp. 111-156, IRL Press at Oxford University Press, Oxford, U.K.). Salt concentration, melting temperature, the absence or presence of denaturants, and the type and length of nucleic acid to be hybridized (*e.g.*, DNA, RNA, PNA) are some of the variables considered when adjusting the stringency of a particular solution hybridization reaction according to methods known in the art (Ausubel *et al.* 1987-1994, p. 2.10.7-2.10.16; Britten and Davidson, 1985, Hybridization Strategy, In B.D. Hanes and S.J. Higgins, eds., Nucleic Acid Hybridization: A Practical Approach, p. 3-46 IRL Press, Oxford University Press, Oxford; Mullis *et al.*, 1986, Cold Spring Harbor Symp. Quant. Biol. 51:263-73).

In the second step, the annealed hybrids are captured on a solid phase surface. In one embodiment, where a single probe is used, the solid phase surface can be the wells of a microtiter dish. In an exemplary, non-limiting example, the wells of a microtiter plate (*e.g.* Maxisorb) are coated with 100 μ l of an anti-peptide antibody diluted to

a concentration of 4 mg/ml in 0.1 M Na₂CO₃-NaHCO₃ (pH 9.1) for 1.5 hours at 37°C. The plate may then be stored overnight at 4°C. Each well is washed 4 times with 200 µl phosphate-buffered saline (PBS); 0.1% Tween 20 (Sigma); 10 mM EDTA. The wells are then incubated for 2 hours in 100 µl PBS; 3% bovine serum albumin (BSA, Sigma) at room temperature to block non-specific binding sites, and washed again as previously described.

5 The reannealed mixture is brought to a final volume of 120 µl with PBS, transferred to the antibody coated well and incubated at 37°C for 45 minutes. The wells are then washed thoroughly (3 to 4 times) in 200 µl PBS; 0.1% Tween 20; 10 mM EDTA to remove unbound nucleic acids, such as all target molecules that have not hybridized to a probe having a peptide-label.

10 In another embodiment, where multiple probes are captured and analyzed together in batch or "multiplex", the solid phase surface comprises a plurality of loci, each locus comprising an antibody with specific affinity for a peptide linked to an SNP-specific probe (*i.e.* an antibody array chip). A detailed description of such an array and methods for its construction is provided in Section 5.5. The chip can be contacted with one or more
15 SNP-specific probes as described for the microtiter plate above.

 In the third step, the solid phase surface is treated with a single-stranded nuclease, such as mung bean nuclease, to cleave the single-stranded regions on hybrid molecules and to destroy all non-hybridized probes. Mung bean nuclease is a single-strand specific DNA and RNA endonuclease which removes single-stranded extensions from the
20 ends of DNA and RNA duplex molecules, leaving blunt ends, and leaving internal mismatches shorter than three consecutive bases in double stranded DNA:DNA or DNA:RNA hybrids intact. By way of an example, the wells of a microtiter plate can be filled with 150 µl of buffer containing 0.05 units of mung bean nuclease and incubated at
25 37°C for 15 minutes; the wells are then washed 4 times with 200 µl phosphate buffered saline (PBS; 150 mM NaCl-10 mM Na-PO₄, pH 7.4). After this treatment, all target:probe hybrids will contain both the 5' and 3' markers irrespective of the presence or absence of SNP mismatches. Only in those hybrids which contain mismatches greater than three consecutive bases will the 5' and 3' markers be separated from each other.

 The chip is rinsed thoroughly as in step 2, and, in step 5, signals are detected
30 from the markers of an SNP oligonucleotide probe. If the markers were not directly labeled, they can be labeled at this point with any affinity labeling method known in the art. In a preferred embodiment, a partner molecule comprising a detectable label is added to specifically bind an affinity group. For example, the 3' marker of the SNP-specific probe may be linked to a biotin molecule. Streptavidin, a binding partner molecule of biotin, may
35 be linked to an enzyme label, such as horseradish peroxidase (HRP), and used as an affinity binding partner to specifically bind to the biotin-linked marker. In a specific embodiment,

120 μ l of a 1×10^{-5} dilution of Amdex reagent, a streptavidin-linked, agarose-bound, aggregate of about 100 molecules of horseradish peroxidase. After a 30 minute incubation at 25°C, the wells are washed 4 times with 200 μ l PBS to remove excess streptavidin HRP, and 160 μ l of TMB (tetramethyldecanoic acid, a substrate for horse radish peroxidase) solution is added. The resulting colorimetric reaction is allowed to proceed for 5 minutes, 5 60 μ l of liquid is transferred to an uncoated well (of a fresh microtiter plate) to which 25 μ l of 1 M sulfuric acid is added to fix colorimetric development. The intensity of this first reaction is read at 450 nm in a plate reader spectrophotometer. The second marker, the 5' label, may be linked, for example, to a chelate of divalent or trivalent cations (Ni^{2+} , Cu^{2+} , Eu^{3+} , etc.; Dahlen et al, 1994, Bioconjug. Chem. 5:268-72; Li and Selvin, 1997, Bioconjug 10 Chem. 8:127-32). A PDLS-coupled globular protein such as ovalbumin, thyroglobulin etc., may be used to specifically bind the chelate marker (Lamtire and Wensel, 1995, Bioconjug Chem. 6:88-92). PLDS, a polymer of lysine, dipicolinate, and succinate, binds divalent and trivalent cations with high affinity and strongly enhances the visible luminescence they emit when excited with ultraviolet light (Lamtire and Wensel, 1995, Bioconjug. Chem. 6:88-92).

15 In another embodiment, a primary antibody comprising a detectable label is added to specifically bind one or both of the first and second markers. In yet another embodiment, a primary antibody may be added, washed, and followed by addition of a secondary antibody labeled with a detectable label. Such labeling and detection methods are well known in the art, and are further described in Sections 5.4.1 and 5.4.2.

20 The first "baseline" reading is taken to measure the probe's 3' and 5' markers. The signal is normalized by comparing the reading obtained to the reading obtained from a "control" sample. In such a control sample, an oligonucleotide having the sequence of the non-polymorphic site is hybridized to a sample having the non-polymorphic sequence. This reading serves as a measure of the amount of starting material present 25 before the mismatch detection analysis. Here, the probes that produce both signals are hybridized to homologous DNA targets, and probes that produce only 5' signal are hybridized to targets that contain a deletion or an insertion, spanning at least three consecutive bases, relative to the probe sequence. This reading also serves as a baseline reading to normalize samples for differences between various SNP-specific probes in 30 antibody binding affinities.

In the sixth step, the hybrids are digested with one or more enzymes to cleave at mismatched base-pairs. In the preferred embodiment, the deoxyinosine 3' endonuclease E. coli endonuclease V (endoV) and S1 nuclease (S1) are used. EndoV, the product of the E.coli *nfi* gene (also known as dI Endo, GenBank accession number 35 U00006), is a deoxyinosine 3' endonuclease that recognizes and nicks DNA containing deoxyinosine residues, single base mismatches, insertion/deletion mismatches, apurinic

sites, flaps, and pseudo-Y structures (Dempfle and Linn, 1982, J. Biol. Chem. 257:2848-55; Yao and Kow, 1996, J. Biol. Chem. 271:30672-6; Yao and Kow, 1997, J. Biol. Chem. 272:30774-79). In this embodiment, endo V will cleave a mismatched hybrid at the phosphodiester bond of the target strand (non-probe) 3' to the mismatch, either immediately adjacent to or one nucleotide removed, depending on the sequence context of the mismatched base pair. S1 nuclease cleaves across from the single-stranded nick introduced by endo V, producing a double stranded break. Thus, when a SNP is present, cleavage by Endo V and S1 nuclease dissociates the 3' marker from the 5' marker. When a SNP is not present, the 3' and 5' markers remain associated with the same molecule.

In one embodiment, the wells of the microtiter dish are filled with 120 μ l of a reaction buffer containing 0.001 units of S1 nuclease and 0.05 units of endonuclease V. The reaction is allowed to proceed for 15 minutes at 37°C and the wells are then washed as described previously. In an alternative embodiment, where the method is being performed in multiplex, the antibody chip containing the trapped hybrids is placed in a 37°C reaction bath containing 0.01 units/ml of S1 nuclease and 0.5 units/ml of endo V, and allowed to incubate 12-15 minutes. The washes are accomplished by immersing in a bath of PBS.

In step seven, the signal is developed again, and a second reading is taken. In one embodiment, for example, 160 μ l of TMB solution is added and the resulting colorimetric reaction is allowed to proceed for 5 minutes. 40 μ l of 1 M sulphuric acid is added and the intensity of the colorimetric development in this second reaction is read as before. In the alternative embodiment, the antibody array is re-developed and the signal recorded. Array coordinates where the intensity of the 3' signal remains unchanged contain hybridized genomic sequence identical to the probes; coordinates where the 3' signal has disappeared but the 5' signal intensity is unchanged contain hybridized genomic sequence non-identical to the probes, and coordinates where the 3' signal intensity has fallen (e.g. by more than 35%) contain hybridized genomic fragments that are heterozygous with respect to the probes. In another embodiment, coordinates where the 3' signal intensity has fallen by more than 25% contain hybridized genomic fragments that are heterozygous with respect to the probes. In another embodiment, coordinates where the 3' signal intensity has fallen by more than 20% contain hybridized genomic fragments that are heterozygous with respect to the probes. The example in Section 6 demonstrates how one analyzes the signals to detect the presence of a polymorphic site.

In various embodiments, the method is used for multiplex SNP screening. As used herein "multiplex" means simultaneous screening for the presence of more than one SNP site in a single nucleic acid target or a single SNP site in more than one target nucleic acid. In one embodiment, for example, the SNP detection method is used to screen a nucleic acid sample from a single subject for variations across a spectrum of SNP sites.

This embodiment of the method may be useful for various purposes, including diagnosing illness, determining a subject's predisposition to genetic disorders, determining a patient's tolerance or toxicity to a drug treatment, and for identifying or determining the ancestry of a subject. In this embodiment, a library of SNP-specific peptide-labeled oligonucleotides is used to probe a single DNA sample for the presence of SNPs. A library of SNP-specific peptide-labeled oligonucleotides, comprising pairs of "positive" probes and "negative" probes of known SNP sequences, may be used to screen the nucleic acid of a single patient. Hybridizations may be performed in batch, annealing many probes to target nucleic acid in a single mixture. The product is analyzed on addressable antibody arrays, each probe identifiable by its array "address", or x-y coordinates.

In another embodiment, the SNP detection method can be used for multiplex screening in a phenotype/genotype association analysis. In this embodiment, a target DNA sample comprises a plurality of DNA molecules derived from a population of individuals afflicted with a disorder of interest (for example, arthritis, asthma, or cancer). A control DNA sample comprises DNA molecules derived from a non-afflicted population. The two populations can be screened in multiplex with a number of SNP probes, and the results compared between the two populations.

In yet another embodiment, the SNP detection method can be used for multiplex screening of a number of individuals for one or more SNP sites. This method is useful, for example, for diagnostic screening of large numbers of patients for a genetic disorder, for drug sensitivity, for blood type, etc. In this embodiment, a SNP-specific probe corresponding to a specific SNP may be linked with any number of different peptide labels. Each SNP probe is annealed to a different sample of target nucleic acid separately, so that each sample is differentially labeled. The samples may then be processed together "in batch" and detected and analyzed using addressable antibody arrays.

The addressable array chip can be regenerated and used repeatedly. For example, peptides can be released from antibodies on an addressable array chip by using a pH gradient which causes the antibodies to release the trapped peptides. The chip is then thoroughly rinsed in PBS for storage or immediate reuse.

5.3 METHODS FOR ANALYSIS OF GENE EXPRESSION USING ADDRESSABLE ANTIBODY ARRAYS

In another embodiment of the invention, methods are provided to monitor gene expression events, both qualitatively and quantitatively, using peptide-labeled oligonucleotide probes and addressable antibody arrays. The methods described herein can be used, for example, to detect or measure the level or abundance of a wild type or polymorphic RNA, including an alternatively spliced or mis-spliced RNA, in a tissue sample, or to directly monitor qualitative and quantitative gene expression levels in tissue

biopsies or in histological preparations. Such biopsies and histological preparations may be taken from a patient with a suspected genetic disorder, for example, to identify the absence or presence of one or more mutations therein. Peptide-labeled oligonucleotide probes further labeled with additional markers (*e.g.* fluorescent labels) may be used to monitor such known sites of potential genetic variation, particularly at RNA splice junctions.

- 5 Particularly interesting targets are known mutations occurring at intron-exon boundaries that lead to aberrantly-spliced mRNAs and result in a broad spectrum of genetic disorders. Mutations commonly occur at such splice-site junction sequences, the region of mRNA spanning the boundary of two exons, formed by a splicing event that removes an intervening sequence of non-coding RNA from a precursor mRNA molecule.
- 10 Indeed, fifteen percent of all point mutations that result in human genetic disease create an RNA splicing defect (Krawczak *et al.*, 1992, Hum. Genet. 90:41-54). Such splicing mutations frequently affect exon sequences near the intron/exon borders; many of these splice-site mutations are "silent" mutations, resulting in no alteration in the primary amino acid sequence of the encoded protein. Some non-limiting examples include: a G→A
- 15 mutation in the integrin GPIIIa gene resulting in thrombocytopenia (Jin *et al.*, 1996, J. Clin. Invest. 98:1745-54); a C→T mutation in the gene for fumarylacetoacetate hydrolase resulting in hereditary tyrosinemia (van Amstel *et al.*, 1996, Hum. Genet. 97:51-59); an A→G mutation in pyruvate dehydrogenase E1a resulting in encephalomyelopathy (de Meirleir *et al.*, 1994, Pediatr. Res. 36:707-712); and numerous others (see generally, Cooper and Mattox, 1997, Am. J. Hum. Genet. 61:259-266).
- 20

- A splice-site probe is designed to span a splice-site junction sequence of a given RNA where either a genomic splice site mutation or an alternative splicing event is known or suspected to occur. A diagrammatic representation of such a mutation, and the splice-site oligonucleotide probe that spans the splice junction between exon 1 and exon 2, is shown in Figure 3.
- 25

- Total ribonucleic acid (RNA) may be isolated from a tissue sample by a variety of methods known in the art, depending on the source and amount of cells available. Purified RNA can be prepared from tissue samples recovered from a biopsy, from tissue culture cells, etc. Procedures described in standard treatises (*e.g.*, Methods in Enzymology, 30 1987, volume 154, Academic Press; Sambrook *et al.*, 1989, *supra*; and Ausubel *et al.*, *supra*) may be followed to carry out routine molecular biology reactions in purification of total nuclear and cytoplasmic RNA. Methods described in detail *infra* are for illustration only and not by way of limitation. Various RNA preparation systems that are commercially available may also be used according to the manufacturer's instructions for isolating and
- 35 purifying the RNA to be used in the methods of the invention. It is preferable to obtain high quality RNA that is of high molecular weight in order to recover RNA containing even rarely expressed gene products. To prepare quality RNA, methods that provide complete

lysis of cells and rapid inactivation of nucleases are preferred. Many such methods are known in the art. For example, one single-step RNA preparation method uses the strong chaotropic agent, guanidinium isothiocyanate, together with a mild detergent and β -mercaptoethanol or dithiothreitol, to denature proteins and inactivate nucleases, and purification of the RNA by ultracentrifugation (Chomczynski & Sacchi, 1987, Anal. Biochem. 162:156-159; Chomczynski, 1989, U.S. Patent No. 4,843,155). This method may also be used especially when isolating RNA from small quantities of cellular material.

The purified RNA is next heated to 65°C in hybridization buffer (e.g., 10-50 μ g RNA in 150 mM KCl, 10 mM TrisCl pH 8.3, 1 mM EDTA) to destroy all secondary structures, and an aliquot of the entire collection of probes is added to the mixture. The RNA is then mixed with the probes, and slow-cooled to room temperature to allow hybridization to occur. In the preferred embodiment, to identify and measure RNA transcripts, the SNP detection method described in Section 5.2 is followed, substituting SNP-specific probes with splice-site probes and using RNA target nucleic acids instead of DNA target nucleic acids. After the hybridization reaction, the sample is contacted with a solid phase surface comprising a binding partner to the peptide label of the oligonucleotide. In various embodiments, for example, the solid phase surface may be a microtiter plate or, for purposes of multiplex gene expression analysis, an addressable antibody array (e.g., on a plastic chip).

The solid phase is then treated with mung bean nuclease, which cleaves single stranded regions (and any excess probe) but does not recognize or cleave mismatched base pairs. An initial reading is taken (R_1 in Figure 3). Where the two markers present on the probe are fluorophores (e.g. F_1 and F_2 in Figure 3), their emission wavelengths are monitored to identify any addressable loci which retain probes that emit at one or both wavelengths. The initial reading also provides an internal baseline, representing the number of RNA molecules captured by the probes. Accordingly, by using an excess of probe relative to RNA, it is possible to quantitate precisely the amount of a specific transcript in a sample being analyzed, by quantitating the readout from the initial reading.

Where a splice event that would result in the formation of the splice-site junction corresponding to a given probe has not occurred in a given sample, the mung bean nuclease treatment results in partial digestion of the probe and the release of the 3' probe fragment, comprising the F_2 fluorophore. Thus, where the RNA in the sample was incorrectly spliced or alternatively spliced at the splice junction corresponding to the sequence of the splice-site probe, the probe remaining bound to the solid support after the mung bean nuclease treatment retains the F_1 fluorophore but loses the F_2 fluorophore to mung bean nuclease digestion (see example in Figure 3). An alternatively spliced transcript is thereby identified. A sample expressing such RNAs is a candidate for possessing a splice site mutation at the genomic DNA level. The sample is then treated with *E. coli*

endonuclease V (endo V) plus S1 nuclease (S1) and a second reading is taken (R_2 in Figure 3). Analysis of the readout is carried out as described in Section 5.2.

In an alternative embodiment, excess fluorescently-labeled "empty" (*i.e.* un-annealed) probe in the hybridization mixture may be eliminated after the annealing step, to reduce their contribution to the signal in the final product. This may be accomplished, for example, by passing the mixture of RNA and probes over a hydroxyapatite column. Single-stranded probes are trapped in the column, while annealed probes bound to RNA appear in the column flow-through. In another embodiment, empty probe may be separated from hybridized probe by electrophoresis. Such electrophoresis may be accomplished, for example, by a lateral electrophoresis step carried out within a chamber connecting the hybridization vessel and the chip. In another embodiment, an "anti-probe" may be used to eliminate the fluorescent signal from excess un-annealed probe. The anti-probe comprises an oligonucleotide complementary to the peptide-labeled probe that has two fluorescent quenchers matched to the fluorophores present on the probe of interest. An anti-probe can be added and annealed to the reaction at an appropriate step in order to "mop up" fluorescence which would otherwise result from the presence of excess un-annealed probe. Accordingly, various alternative embodiments are available for practicing the invention, as exemplified above, which may be selected by the practitioner to suit various needs.

5.4 METHODS FOR LABELING AND DETECTION OF THE OLIGONUCLEOTIDE PROBE

Various labeling, signal amplification, and detection methods are important for increasing the sensitivity of the PCR-free high-throughput methods of the invention. Signal amplification and enhanced detection methods allow greater detection of nucleic acid sequences present in low abundance and/or from small amounts of tissue sample. Such methods also facilitate multiplex screening of larger sample sizes in a single run. Methods for labeling probes with detectable labels, amplifying and detecting a signal, preparing oligonucleotides, and linking oligonucleotides to peptides are described herein below.

5.4.1 LABELING METHODS

A "label" or a "detectable label" refers to a moiety, such as a luminescent, fluorescent, or radioactive isotope or group containing same, and nonisotopic labels, such as enzymes or dyes. For the embodiments presented herein, where two or more markers are present on a single probe, it is important that the two detectable labels are distinguishable. This may be accomplished, for example, by using different fluorophores that emit at different frequencies, or by using affinity partner molecules containing different fluorophores, as discussed in Section 5.4.2, or by using different types of luminescent agents, enzymes, or dyes.

The oligonucleotide probe may be labeled with a detectable label either directly or indirectly. In a preferred embodiment, the SNP-specific oligonucleotide probe is labeled indirectly by incorporating an affinity group into the oligonucleotide probe during its synthesis (for methods for synthesizing oligonucleotides, see Section 5.4.3, below). A partner molecule specifically recognizes and binds to the affinity group. The partner molecule can comprise a detectable label, such as a fluorophore, coupled to the affinity binding partner. Alternatively, a second partner molecule having affinity for the first partner molecule (such as a secondary antibody) can be used. In this case the second, and not the first, affinity binding partner will comprise the detectable label. Indirect affinity labeling methods are well known in the art (see, generally, Current Protocols in Immunology, Coligan *et al.*, eds., 1997, John Wiley & Sons, Inc. U.S.A., pp. 8.10.12 - 8.10.21).

In a preferred embodiment, a peptide-labeled oligonucleotide is labeled with biotinylated nucleotides and detected using a streptavidin-linked detectable compound. Biotinylated oligonucleotides are well known in the art. For example, an oligonucleotide may be biotinylated using a biotin-NHS ester procedure. Alternatively, biotin may be attached during oligonucleotide synthesis using a biotin phosphoramidite (Cocuzza, 1989, Tetrahed. Lett. 30, 6287-6290). One such biotin phosphoramidite available from Glen Research is 1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite. This compound also has a branch point to allow further additions. The branched spacer used in this biotin phosphoramidite has been described by Nelson *et al.* (1992, Nucl. Acids Res. 20, 6253-6259). Another 5'-biotin phosphoramidite, namely [1-N-(4,4'-Dimethoxytrityl)-biotinyl-6-aminohexyl]-2-cyanoethyl-(N,N-diisopropyl)-phosphoramidite, may be used to biotinylate an oligonucleotide. This compound is sold by Glen Research under license from Zeneca PLC.

In another embodiment, for example, the affinity group is a chelate of divalent or trivalent cations (Ni^{2+} , Cu^{2+} , Eu^{3+} , etc.), and a PDLs-coupled globular protein such as ovalbumin, thyroglobulin etc., is used as an affinity binding partner (Dahlen *et al.*, 1994, Bioconjug. Chem. 5:268-72; Li and Selvin, 1997, Bioconjug. Chem. 8:127-32; Lamture and Wensel, 1995, Bioconjug. Chem. 6:88-92).

In another embodiment, the oligonucleotide probe is labeled by incorporation of hapten-labeled nucleotides such as digoxigenin-labeled nucleotides (Holtke and Kessler 1990), and the hapten-labeled bases are detected using an antibody that specifically binds to the hapten, for example anti-DIG-alkaline phosphatase or POD conjugate (Fab fragments). Such compounds can be detected using chemiluminescent, colorimetric, or fluorescent substrates. Some of the affinity labels and reagents for their detection and isolation are available commercially.

In an alternative embodiment, the SNP-specific probe is directly labeled. The detectable labels may be distinct fluorophores (F_1 and F_2) linked to deoxyribose bases, as depicted in Figure 2. Direct methods for linking detectable labels to nucleic acids are well known to those of skill in the art. For example, European Patent Publication No. EP 0370 694 A2, entitled, "Diagnostic Kit and Method Using a Solid Phase Capture Means For
5 Detecting Nucleic Acid", by Burdick and Oakes, publication date May 30, 1990, discloses methods of linking such labels to oligonucleotides. For example, fluorophores can be linked to deoxyribonucleotide bases prior to synthesis of the oligonucleotide primer. Direct labeling can be accomplished by incorporating fluorescent dye-labeled phosphoramidites into an oligonucleotide during synthesis. Fluorescently labeled nucleoside analogs are
10 available commercially, or can be produced synthetically, by methods known in the art (see, Brumbaugh *et al.*, 1988, Proc. Natl. Acad. Sci. (USA), 85:5610-5614). Such analogs can be synthesized by derivatizing a 2'-deoxynucleoside through organometallic intermediates to give the 5-dimethoxytrityl adduct. The methyl ester is hydrolyzed, activated, and reacted with an appropriately monoacylated alkyl diamine. Such linker arm nucleosides can be
15 purified, and converted to nucleoside analogs suitable for chemical oligonucleotide synthesis. Two such fluorescein labels are 5'-hexachloro chloro-fluorescein phosphoramidite (HEX), and 5'-tetra chloro-fluorescein phosphoramidite (TET), both available from Glen Research. Oligonucleotides may be synthesized using any method known in the art (see Section 5.4.3). For example, standard phosphoramidite chemistry
20 may be used and reagents may be obtained from any one of many commercial suppliers (e.g., Applied Biosystems 392/394 DNA synthesizer).

5.4.2 SIGNAL AMPLIFICATION AND DETECTION METHODS

For the embodiments presented herein, where two or more markers are
25 present on a single probe, it is important that the two detectable labels are distinguishable. This may be accomplished, for example, by using different fluorophores that emit at different frequencies, as discussed hereinbelow, by using different haptens or affinity agents, or by using different types of luminescent agents.

In a preferred embodiment, the detectable label is a luminescent agent.
30 Luminescent agents, depending upon the source of exciting energy, can be classified as radioluminescent, chemiluminescent, bioluminescent, and photoluminescent (including fluorescent and phosphorescent). Various fluorescent dyes known in the art may be used. Two commonly used groups of dyes are the fluorescein-based dyes and the rhodamine-based dyes. Specific, non-limiting examples of fluorophores which can be used are given in
35 Table 1.

Table 1
Emission and Absorption Spectra of Commonly Used Fluorophores

	DYE	Absorbance (nm)	Emission (nm)
5	fluorescein	494	525
	6-Pam	494	525
	Joe	520	548
	Tet	521	536
	Hex	535	536
	Cy3	552	570
10	Tamra	565	580
	Rox	568	595
	Texas red	587	602
	Cy5	643	667

In general, fluorescein-based fluorophores tend to absorb short wavelengths (less than 500 nm) and emit photons that have wavelengths approximately 30 nm longer than the absorbed photons, while rhodamine-based fluorophores absorb in the green/yellow range and emit at energies close to that of the absorbed photons (less than 25 nm difference). When using more than one fluorophore, care must be taken in choosing compatible fluorophores to avoid cross-quenching, the result of the emission wavelength of one fluorophore being in the range of the absorption spectra of another fluorophore. In that case, only the emission of the second fluorophore will be detected. Thus, in the most preferred embodiment, the labels are two fluorophores that absorb the same wavelength but emit at different frequencies.

Any method known in the art may be used for visualization or detection of a signal from the detectable label. In one embodiment, the detectable label can be measured by a spectroscopic device. For example, fluorescence can be measured by a device, such as a fluorimeter, or a flow cytometer. Methods for performing assays on fluorescent materials are well known in the art and are described in, *e.g.*, Lackowicz, J.R., 1983, Principles of Fluorescence Spectroscopy, New York:Plenum Press. In another embodiment, the detectable label may be detected by confocal microscopy.

5.4.3 OLIGONUCLEOTIDES AND OLIGONUCLEOTIDE ANALOGS

The oligonucleotide probes used in conjunction with the methods of the invention often comprise oligonucleotides ranging from about 10 to about 100 nucleotides in length. In specific aspects, an oligonucleotide may be 10 nucleotides, 15 nucleotides, 25 nucleotides, 50 nucleotides, 100 nucleotides, or 200 nucleotides in length, or any number of

nucleotides ranging from 10 to 500 nucleotides in length. In preferred embodiments, an oligonucleotide probe is from 35 to 65 nucleotides in length, or approximately 50 nucleotides in length. An oligonucleotide can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, or single-stranded or double-stranded, or partially double-stranded. An oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, or a combination thereof. An oligonucleotide may include other
5 appending groups, such as biotin, fluorophores, or peptides. Where modified oligonucleotides are used, consideration should be given to the effect on enzyme activity (e.g., S1 nuclease, endo V, etc.) such modifications may have.

An oligonucleotide or derivative thereof used in conjunction with the
10 methods of this invention may be synthesized using any method known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein *et al.* (1988, Nucl. Acids Res. 16, 3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass
15 polymer supports (Sarin *et al.*, 1988, Proc. Nat'l Acad. Sci. U.S.A. 85, 7448-7451), etc. An oligonucleotide may be an α -anomeric oligonucleotide. An α -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (see Gautier *et al.*, 1987, Nucl. Acids Res. 15, 6625-6641).

Oligonucleotides may be synthesized using any method known in the art
20 (e.g., standard phosphoramidite chemistry on an Applied Biosystems 392/394 DNA synthesizer). Further, reagents for synthesis may be obtained from any one of many commercial suppliers. Spacer phosphoramidite molecules may be used during oligonucleotide synthesis, e.g., to bridge sections of oligonucleotides where base pairing is
25 undesired or to position labels or tags away from an oligonucleotide portion undergoing base pairing. The spacer length can be varied by consecutive additions of spacer phosphoramidites. Spacer phosphoramidite molecules may be used as 5'- or 3'-oligonucleotide modifiers. Such spacers include Spacer Phosphoramidite 9 (i.e., 9-O-Dimethoxytrityl-triethyleneglycol, 1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite, and Spacer Phosphoramidite 18 (i.e., 18-O-Dimethoxytrityl-hexaethyleneglycol, 1-[(2-
30 cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite), both available from Glen Research (Sterling, Virginia).

Other spacers are available for use in standard oligonucleotide synthesis. For
example, Spacer Phosphoramidite C3 and dSpacer Phosphoramidite can be used to
35 destabilize undesirable self-hybridization events within oligonucleotide probes or to destabilize false hybridization events between incorrectly-matched template/probe complexes. This spacer is available from Glen Research, (i.e., 3-O-Dimethoxytrityl-propyl-

1-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite), and can be added to substitute for an unknown base within an oligonucleotide sequence.

A branching spacer may be used as one method to increase label incorporation into an oligonucleotide. Such a branching spacer may also be used to increase a detectable signal by hybridization through multiply branched probes. Branching spacers are also available commercially, *e.g.*, from Glen Research.

Biotinylated oligonucleotides are well known in the art. An oligonucleotide may be biotinylated using a biotin-NHS ester procedure. Alternatively, biotin may be attached during oligonucleotide synthesis using a biotin phosphoramidite (Cocuzza, 1989, *Tetrahed. Lett.* 30, 6287-6290). One such biotin phosphoramidite available from Glen Research is 1-Dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite. This compound also has a branch point to allow further additions. The branched spacer used in this biotin phosphoramidite has been described by Nelson *et al.* (1992, *Nucl. Acids Res.* 20, 6253-6259).

Another 5'-biotin phosphoramidite, namely [1-N-(4,4'-Dimethoxytrityl)-biotinyl-6-aminohexyl]-2-cyanoethyl-(N,N-diisopropyl)-phosphoramidite, may be used to biotinylate an oligonucleotide. This compound is sold by Glen Research under license from Zeneca PLC.

An oligonucleotide may comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, and 2,6-diaminopurine.

An oligonucleotide may comprise at least one modified phosphate backbone selected from the group including but not limited to a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

5.4.4 LINKING OLIGONUCLEOTIDES WITH PEPTIDES

Labeled oligonucleotides have long been used for detection of specific sequences. For example, Burdick and Oakes (Diagnostic kit and method using a solid phase capture means for detecting nucleic acids, European Patent Publication No. EP 0370 694 A2, published May 30, 1990) disclose the use of labeled oligonucleotides, with specific nucleic acid sequences which are complementary to a predetermined sequence of interest. Methods for manipulating nucleic acids on antibody solid phase supports using peptide-labeled oligonucleotides have been described previously (see pending U.S. Patent application, Application No. 09/174,328, by Iris and Pourny, filed September 16, 1998, which is incorporated herein by reference in its entirety).

In a preferred embodiment, a peptide label is linked to an oligonucleotide probe. The peptide label is used as an affinity label for binding to chip-based antibody arrays. Methods of attaching peptides to oligonucleotides are well known to those with ordinary skill in the art, *see, e.g.*, Soukchareun S. *et al.*, 1995, *Bioconjug. Chem.* 6:43-53; Tung CH *et al.*, 1991, *Bioconjug. Chem.* 2:464-465; Bruick RK *et al.*, 1996, *Chem. Biol.* 3:49-56; Tung CH *et al.*, 1995, *Bioconjug. Chem.* 6:292-295; Robles J. *et al.*, *Bioconjug. Chem.*, 1997, 8:785-788; and Rajur S.B., *et al.*, 1997, *Bioconjug. Chem.* 8:935-940.

Oligonucleotides linked to various peptides for use in the methods of this invention may be obtained, for example, from Cybergene S.A. (11 rue Claude Bernard, z1 nord, 35400, Saint Mallo, France) and Glen Research (22825 Davis Drive, Sterling, Virginia 20164). Further information from Glen Research can be obtained through their web site (<http://www.glenres.com>).

One specific method for linking a peptide to an oligonucleotide recommended by Glen Research is as follows (*see also*, <http://www.glenres.com>). A heterobifunctional crosslinking reagent is used to link a synthetic peptide having an N-terminal lysine residue to a 5'-thiol-modified oligonucleotide. Such a crosslinking reagent is N-maleimido-6-aminocaproyl-(2'-nitro, 4'-sulfonic acid) phenyl ester (mal-sac-HNSA). The sodium salt of mal-sac-HNSA is available from Bachem Bioscience. Conveniently, reaction of the mal-sac-HNSA crosslinker with an amino group releases a dianion phenolate (*i.e.* 1-hydroxy-2-nitro-4-benzene sulfonic acid). This dianion phenolate is also a yellow chromophore. The chromophore feature provides (i) a means for quantifying the extent of completion of the coupling reaction (where greater yellow color intensity corresponds to a more complete coupling reaction), and (ii) an aid in monitoring the extent of separation of an activated peptide (*i.e.* a peptide crosslinked to mal-sac-HNSA and ready for contacting with a 5'-thiol-modified oligonucleotide) from free crosslinking reagent during gel filtration.

The specific steps employed when using a mal-sac-HNSA crosslinker may be as follows. First, a peptide is synthesized having an N-terminal lysine. Alternatively, a peptide having an internal lysine may be used since the lysine epsilon amino group is

actually more reactive than the lysine alpha amino group. Second, an oligonucleotide is synthesized having a 5'-thiol group using methods known in the art. Third, the peptide is reacted with an excess of mal-sac-HNSA in a sodium phosphate buffer (pH 7.1). Fourth, the peptide-mal-sac conjugate is separated from free crosslinker and the buffer is exchanged to sodium phosphate (pH 6) using a gel filtration column (*e.g.* NAP-5, Pharmacia, Uppsala, Sweden). Fifth, a thiol-modified oligonucleotide is activated, desalted and buffer-exchanged to sodium phosphate (pH 6) on a gel filtration column. Sixth, the activated peptide is reacted with the thiol-modified oligonucleotide. Finally, the peptide-oligonucleotide conjugate is purified by ion exchange chromatography (*e.g.* Nucleogen DEAE-500-10 or equivalent). The elution order from the ion exchange column is as follows: free peptide first, peptide-labeled oligonucleotide next, and free oligonucleotide last.

In a preferred embodiment, a two-part (*i.e.* of two distinct molecules) peptide-labeled oligonucleotide probe may be constructed for use in the methods of the invention (*see* Figure 4). A two-molecule probe may be constructed from a first molecule comprising nucleotides of any length attached to a peptide label (*i.e.* a peptide-oligonucleotide conjugate manufactured or obtained as described above), and a second molecule comprising nucleotides complementary to the first molecule contiguous with additional nucleotides specific for detection of a given polymorphism (*i.e.* complementary to either a wild type or variant form of the target polymorphism and its flanking regions). In a preferred embodiment, the two-molecule probe is constructed from a first molecule comprising at least 15 nucleotides attached to a peptide label, and a second molecule comprising at least 15 nucleotides complementary to the first molecule contiguous with from 30 to 40 nucleotides specific for detection of a given SNP. In another preferred embodiment, the two-molecule probe is constructed from a first molecule, the nucleotide sequence of which consists of from 15 to 25 nucleotides attached to a peptide label; and a second molecule, the nucleotide sequence of which consists of from 15 to 25 nucleotides complementary to the first molecule and contiguous with from 30 to 40 nucleotides specific for detection of a given SNP, wherein at least one nucleotide of the second molecule is labeled with a detectable label. The two molecules or parts are then assembled by hybridization prior to use for detection of a given SNP (Figure 4). An advantage of employing a two-part probe is the significant cost savings associated with having a set of standardized peptide-oligonucleotide conjugates, which are relatively expensive to produce, which are then made specific for detection of various SNP's of interest simply by synthesizing the second part of the probe, which is relatively inexpensive to produce.

When employing a two-part probe, it may be desirable to include standard nuclease-resistant nucleotide base derivatives (*e.g.* thiolated bases) at positions near the transition from double-stranded to single-stranded structure within the assembled probe (*see*

shaded rectangles in Figure 4). Accordingly, in a preferred embodiment, three thiolated bases may be placed at the 3' end of a peptide-oligonucleotide conjugate, wherein the peptide label is attached at the 5' end, and an additional three thiolated bases may be placed such that they occur as the first three bases in the single-stranded portion of the assembled probe immediately adjacent to the double-stranded portion (*see* shaded rectangles in Figure 4). In a further preferred but non-limiting aspect, the first or second base immediately adjacent to the thiolated bases in the second part of the probe may be labeled with a first label (*e.g.* F₁ in Figure 4) and the last base at the free single-stranded end may be labeled with a second label (*e.g.* F₂ in Figure 4).

This invention further provides a kit comprising in one or more containers a purified first molecule comprising a first sequence of nucleotides, which molecule is covalently attached to a peptide label, and a purified second molecule comprising a second sequence of nucleotides complementary to the first sequence and contiguous with additional nucleotides complementary to a wild-type or variant form of a nucleotide polymorphism of interest and its flanking regions.

5.5 ADDRESSABLE ANTIBODY ARRAYS

An addressable array, as used in the invention, comprises a plurality of distinct polypeptides attached to precise locations on a solid phase surface, such as a plastic chip. The position of each distinct polypeptide on the surface is known and therefore "addressable". In the preferred embodiment, the polypeptides are distinct antibodies that have specific affinity for the a peptide-linked oligonucleotide probe.

The solid phase surface may be of any material, including, but not limited to, plastic, polycarbonate, polystyrene, polypropylene, polyethylene, glass, nitrocellulose, dextran, nylon, metal, and photolithographic materials. The solid phase surface is preferably a chip, but may also comprise an alternative configuration such as beads, membranes, microparticles, the interior surface of a reaction vessel such as a microtiter plate, test tube or other reaction vessel.

In a preferred embodiment, the antibodies are covalently linked to a plastic chip through their Fc domains. In another embodiment, antibodies are adsorbed onto the chip. Methods of constructing protein arrays, including antibody arrays, are known in the art (*see, e.g.*, Fodor *et al.*, 1996, "Photolabile nucleoside and peptide protecting groups", US Patent No. 5,489,678; Barret *et al.*, 1993, "Spatially-addressable immobilization of anti-ligands on surfaces", US Patent No. 5,252,743; Blawas and Reichert, 1998, "Protein patterning", Biomaterials 19:595-609; Blawas *et al.*, 1996, "Patterning antibodies for multiple analyte sensor via photodeprotection chemistry", San Jose: SPIE; Delamarche *et al.*, 1996, "Immobilization of antibodies on a photoactive self-assembled monolayer on gold", Langmuir 12, 1997-2006; Firestone *et al.*, 1996, "Film architecture in biomolecular

assemblies, Effect of linker on the orientation of genetically engineered surface-bound proteins", J. Amer. Chem. Soc. 118, 9033-9041; Mooney *et al.*, 1996, Patterning of functional antibodies and other proteins by photolithography of silane monolayers, Proc. Natl. Acad. Sci. 93, 12287-12291; Pirrung *et al.*, 1996, "A general method for the spatially defined immobilization of biomolecules on glass surfaces using 'caged' biotin", Bioconjugate Chem. 7, 317-321; Gao *et al.*, 1995, "Immunosensing with photoimmobilized immunoreagents on planar optical wave guides", Biosensors Bioelectron 10, 317-328; Schena *et al.*, 1995, "Quantitative monitoring of gene expression patterns with a complementary DNA microarray", Science 270, 467-470; Lom *et al.*, 1993, "A versatile technique for patterning biomolecules onto glass coverslips", J. Neurosci. Methods, 385-397; Pope *et al.*, 1993, "New applications of silane coupling agents for covalently binding antibodies to glass and cellulose solid phase surfaces", Bioconjugate Chem. 4, 116-171; Schramm *et al.*, 1992, "Antibody-antigen complex formation with immobilized immunoglobulins", Anal. Biochem. 205, 47-56; Gombotz *et al.*, 1991, Protein adsorption to poly(ethylene oxide) surfaces, J. Biomed. Mater. Res. 25, 1547-1562; Alarie *et al.*, 1990, "Evaluation of antibody immobilization techniques for fiber optic-based fluoroimmunosensing", Analy. Chim. Acta 229, 169-176; Owaku *et al.*, 1993, Optical immunosensing for IgG, Sensors Actuators B, 13-14, 723-724; Bhatia *et al.*, 1989, "Use of thiol-terminal silanes and heterobifunctional cross linkers for immobilization of antibodies on silica surfaces", Anal. Biochem. 178, 408-413; Lin *et al.*, 1988, "Characterization of immobilized antibodies on silica surfaces", IEEE Trans. Biomed. Engng., 35(6), 466-471).

In a preferred embodiment, the antibodies are arrayed on a "plastic" chip comprised of electronically activated copolymers of a conductive polymers and an antibodies. Such arrays are known in the art and have been used to link polynucleotides to solid supports for polynucleotide based arrays (U.S. Patent No. 5,837,859 issued November 17, 1998; PCT publication WO 94/22889 dated October 13, 1994). The arrayed pattern may be computer generated and stored. The chips may be prepared in advance and stored in PBS at 4°C. The antibody array chip can be regenerated and used repeatedly. For example, peptides can be released by contacting a pH gradient with the chip causing the antibodies to release the trapped peptides. The chip is then thoroughly rinsed in phosphate buffered saline for storage or immediate reuse.

5.5.1 ANTIBODIES AND PEPTIDES

Antibodies of use with the methods of this invention include any antibodies known in the art. Such antibodies are typically used, for example, to manipulate the nucleic acids of interest. In this regard, a nucleic acid may be manipulated by antibody binding to a portion of the nucleic acid itself or to another antigen (*e.g.*, a protein, peptide or hapten) which is bound (either covalently or non-covalently) to the nucleic acid. In a preferred

embodiment, nucleic acids are manipulated using peptide antigens covalently attached to oligonucleotide probes. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric and humanized antibodies, as described below. Further, single chain antibodies, Fab fragments and F(ab')₂ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above may also be used.

5 Polyclonal antibodies which may be used with the invention are heterogeneous populations of antibody molecules derived from the sera of immunized animals. Various procedures well known in the art may be used for the production of polyclonal antibodies to an antigen-of-interest. For example, the production of polyclonal
10 antibodies, various host animals can be immunized by injection with an antigen of interest or derivative thereof, including but not limited to rabbits, mice, rats, *etc.* Various adjuvants may be used to increase the immunological response, depending on the host species, and including but not limited to Freud's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, polyanions, peptides,
15 oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and corynebacterium parvum. Such adjuvants are also well known in the art.

Monoclonal antibodies which may be used with the invention are homogeneous populations of antibodies to a particular antigen. A monoclonal antibody
20 (mAb) to an antigen-of-interest can be prepared by using any technique known in the art which provides for the production of antibody molecules by continuous cell lines in culture. These include but are not limited to the hybridoma technique originally described by Kohler and Milstein (1975, *Nature* 256: 495-497), the more recent human B cell hybridoma technique (Kozbor *et al.*, 1983, *Immunology Today* 4:72), and the EBV-hybridoma
25 technique (Cole *et al.*, 1985, Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAbs of use in this invention may be cultivated *in vitro* or *in vivo*.

Monoclonal antibodies which may be used with the methods of the invention
30 include but are not limited to human monoclonal antibodies. Human monoclonal antibodies may be made by any of numerous techniques known in the art (*e.g.*, Teng *et al.*, 1983, *Proc. Natl. Acad. Sci. U.S.A.* 80, 7308-7312; Kozbor *et al.*, 1983, *Immunology Today* 4, 72-79; Olsson *et al.*, 1982, *Meth. Enzymol.* 92, 3-16).

A chimeric antibody may be used with the methods of the invention. A
35 chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. Various techniques are available for the

production of such chimeric antibodies (*see, e.g.*, Morrison *et al.*, 1984, Proc. Natl. Acad. Sci. U.S.A. 81:6851-6855; Neuberger *et al.*, 1984, Nature, 312:604-608; Takeda *et al.*, 1985, Nature 314:452-454) by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity.

5 A humanized monoclonal antibody may be used with the methods of the invention. Humanized antibodies are antibody molecules from non-human species having one or more complementarity determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. Briefly, an immunoglobulin light or heavy chain variable region consists of a "framework" region interrupted by three
10 hypervariable regions, referred to as complementarity determining regions (CDRs). The extent of the framework region and CDRs have been precisely defined (*see*, Karat *et al.*, 1983, Sequences of proteins of immunological interest, U.S. Department of Health and Human Services). Various techniques have been developed for the production of humanized antibodies (*see, e.g.*, Queen, U.S. Patent No. 5,585,089; Winter, U.S. Pat. No. 5,225,539, which are incorporated herein by reference in its entirety).

15 Alternatively, techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778; Bird, 1988, Science 242:423-426; Huston *et al.*, 1988, Proc. Natl. Acad. Sci. U.S.A. 85:5879-5883; and Ward *et al.*, 1989, Nature 334:544-546) can be adapted to produce single chain antibodies useful in the device of the invention. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv
20 region together via an amino acid bridge, resulting in a single chain polypeptide.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')₂
25 fragments. Alternatively, Fab expression libraries may be constructed (Huse *et al.*, 1989, Science 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

Selection of antibodies or fragments thereof which may be useful in the methods of the invention may be facilitated by utilization of an antibody or fragment phage
30 display library. Such an approach has the advantage of providing methods for efficiently screening a library having a high complexity (*e.g.* 10⁹), thereby dramatically increasing the chances of identifying antibodies or fragments which are preferred for use in the methods of the invention. Methods for cloning an immunoglobulin repertoire ("repertoire cloning") are well known in the art, can be performed entirely *in vitro*, and permit the production of
35 virtually any kind of antibody without involving an antibody-producing animal. Such procedures are described by way of example briefly below.

Messenger RNA (mRNA) is extracted from human B lymphocytes obtained from peripheral blood. The mRNA serves as a template for cDNA synthesis using reverse transcriptase and standard protocols. Immunoglobulin cDNAs are then specifically amplified by PCR from this very complex mixture of cDNA. In order to construct immunoglobulin fragments with the binding properties of natural antibodies, PCR products from genes encoding antibody light (L) and heavy (H) chains are obtained. The products are then introduced into a phagemid vector. Filamentous bacteriophages (phagemids) are used to infect F+ E. coli. Expression of this phagemid vector, through the intervention of a helper phage such as VCS M13 or M13 K07 in the appropriate bacterial host, results in the expression of an scFv immunoglobulin molecule on the surface of the bacteriophage. The bacteriophage population represents a mixture of immunoglobulins with all specificities included in the repertoire. Antigen-specific immunoglobulin is selected from this population by an iterative process of antigen immunoabsorption followed by phage multiplication. When a bacteriophage specific only for an antigen of interest remains following multiple rounds of selection, it is introduced into a new host which will express a phage-encoded protein in soluble form and in large amounts. The resulting proteins are human antibody fragments with structures very similar to natural antibodies.

Phage-display technology for production of recombinant human antibodies and fragments thereof has been successfully used to produce proteins capable of binding to a broad range of diverse antigens, such as proteins, peptides, nucleic acids, sugars, etc. Further details of this powerful approach may be found in the following references (Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J. Griffiths, A.D., and Winter, G., 1991, "By-passing immunization: human antibodies from V-gene libraries displayed on phage," *J. Mol. Biol.* 222, 581; Nissim, A., Hoogenboom, H.R., Tomlinson, I.M., Flynn, G., Midgley, C., Lane, D., and Winter, G., 1994, "Antibody fragments from a "single pot" phage display library as immunochemical reagents," *EMBO J.* 13, 692; Hoet, R.M.A., F.H.J. van den Hoogen, J.M.H. Raats, P.E.M.H. Litjens, R.M.T. de Wildt, and W.J. van Venrooij, 1996, "Isolation and characterization of antibody fragments directed to disease markers isolated from patient derived combinatorial libraries," In *Keystone Symposium* entitled "Exploring and Exploiting Antibody and Ig Superfamily Combining Sites", Taos, New Mexico, USA, Abstract; Van Kuppevelt, T.H., Dennissen, M.A.B.A., Veerkamp, J.H., Van Venrooij, W.J., and Hoet, R.M.A., 1996, "Anti-heparan sulfate antibody fragments derived from a phage display library," *XVth FECTS Meeting*, Munich, Germany; De Wildt, R.M.T., Finnern, R. Ouwehand, W.H.Griffiths, A.D., Van Venrooy, W.J., and Hoet, R.M.A., 1996, "Characterization of human variable domain antibody fragments against the U1 RNA-associated A protein, selected from a synthetic and patient derived combinatorial V gene library," *Eur. J. Immunol.* 26, 629-639; De Wildt, R.M.T., Steenbakkers, P.G., Pennings, A.H.M., van den Hoogen F.H.J., van Venrooij, W.J., and R.M.A. Hoet, 1997, "A

new method for analysis and production of monoclonal antibody fragments originating from single human B-cells," J. Immunol. Methods. 207, 61-67; Willems, P.M., Hoet, R.M., Huys, E.L., Raats, J.M., Mensink, E.J. and R.A. Raymakers, 1998, "Specific detection of myeloma plasma cells using anti-idiotypic single chain antibody fragments selected from a phage display library," Leukemia 12, 1295-1302; van Kuppevelt, T.H., Dennissen, M.A.B.A., van Venrooij, W.J., Hoet, R.M.A., and J.H. Veerkamp, 1998, "Generation and application of type-specific anti-heparan sulfate antibodies using phage display technology, Further evidence for heparan sulfate heterogeneity in the kidney," J. Biol. Chem. 273, 12960-12966; Hoet, R.M.A., Raats, J.M.H., de Wildt, R., Dumortier, H., Muller, S., van den Hoogen, F., and W.J. van Venrooij, 1998, "Human monoclonal autoantibody fragments from combinatorial antibody libraries directed to the U1snRNP associated U1C protein, epitope mapping, immunolocalization and V-gene usage," Mol. Immunol. 35, 1045-1055).

Further general methods of antibody production and use are suitable for use in connection with the methods of the invention. For example, see Harlow and Lane, 1988, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, which is incorporated herein by reference in its entirety.

The single-letter amino acid code corresponds to the three-letter amino acid code of the Sequence Listing set forth hereinbelow, as follows: A, Ala; R, Arg; N, Asn; D, Asp; B, Asx; C, Cys; Q, Gln; E, Glu; Z, Glx; G, Gly; H, His; I, Ile; L, Leu; K, Lys; M, Met; F, Phe; P, Pro; S, Ser; T, Thr; W, Trp; Y, Tyr; and V, Val.

Suitable antibodies for use with the methods of this invention include the following, available from Affinity Bioreagents, Inc., 79, rue des Morillons, 75015, Paris, France.

- 1) Catalog No. PA 1-047 (affinity-purified rabbit IgG). The corresponding peptide recognized by this Ab is KFSREKKAAKT (SEQ ID NO:1).
- 2) Catalog No. PA 1-039 (affinity-purified rabbit immunoglobins). The corresponding peptide recognized by this Ab is DQKRYHEDIFG (SEQ ID NO:2).
- 3) Catalog No. PA 1-036 (purified rabbit IgG). The corresponding peptide recognized by the Ab is DLKEEKDINNNVKKT (SEQ ID NO:3).
- 4) Catalog No. PA 1-014 (purified rabbit antibody). The corresponding peptide recognized by this Ab is CTGEEDTSE (SEQ ID NO: 4).

- 5) Catalog No. PA 3-013 (affinity purified IgG). The corresponding peptide recognized by this Ab is PEETQTQDQPM (SEQ ID NO:5).
- 6) Catalog No. PA 1-815 (rabbit anti-serum). The corresponding peptide recognized by this Ab is QKSDQGVGPGAT (SEQ ID NO:6).
- 5 7) Catalog No. PA 3-034 (rabbit polyclonal serum IgG). The corresponding peptide recognized by this Ab is DIGQSIKKFSKV (SEQ ID NO:7). This polyclonal antibody will also recognize QRADSLSSHL (SEQ ID NO:8).

10 In addition, antibodies for use with the methods of this invention may be obtained from Medical & Biological Laboratories Co., Ltd., 440 Arsenal Street, Watertown, Massachusetts 02171, U.S.A. These include the following:

- 1) Code No. 561 (Rabbit IgG from anti-serum). The corresponding peptide recognized is YPYDVDPYA (SEQ ID NO:9).
- 15 2) Code No. 562 (Rabbit IgG from anti-serum). The corresponding peptide recognized is EQKLISEEDL (SEQ ID NO:10).
- 3) Code No. 563 (Rabbit IgG from anti-serum). The corresponding peptide recognized is YTDIEMNKLGK (SEQ ID NO:11).
- 20

5.6 TARGET GENETIC DISORDERS

In one embodiment, the high-throughput SNP screening method of the invention is used for diagnosis of genetic disorders in patients' samples. Non-limiting examples of genetic disorders that can be tested using this method include hereditary diseases such as cystic fibrosis, Tay-Sachs disease, sickle cell anemia, hemophilia, atherosclerosis, diabetes, and obesity. Such hereditary diseases may include degenerative and non-degenerative neurological diseases such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, Wilson's disease, spinal cerebellar ataxia, Friedreich's ataxia and other ataxias, prion diseases including Creutzfeldt-Jakob disease, dentatorubral pallidoluisian atrophy, spongiform encephalopathies, myotonic dystrophy, depression, schizophrenia, and epilepsy. Hereditary diseases may also include metabolic diseases such as, for example, hypoglycemia or phenylketonuria. Cardiovascular diseases and conditions are also included, non-limiting examples of which include atherosclerosis, myocardial infarction, and high blood pressure. The invention can further be used for detection and diagnosis of Lyme disease, tuberculosis, and sexually transmitted diseases.

25

30

35

In another embodiment, the SNP detection methods of the invention can be used for determining the genetic basis of a disease or disorder. For example, the SNP detection methods can be used to detect and compare polymorphisms present in patients afflicted with a disorder whose genetic basis is not known. This application of the multiplex PCR-free SNP detection method is particularly useful for complex disorders, where more than one genetic loci may contribute to the phenotype. Once the contributory genetic loci are determined, the SNP detection methods can further be useful in early detection and treatment of such diseases. Non-limiting examples of target disorders of clinical interest include asthma, arthritis, psoriasis, excema, allergies, drug resistance, drug toxicity, and cancers such as, but not limited to, human sarcomas and carcinomas, *e.g.*, fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilms' tumor, cervical cancer, testicular tumor, lung carcinoma, small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, meningioma, melanoma, neuroblastoma, retinoblastoma; leukemias, *e.g.*, acute lymphocytic leukemia and acute myelocytic leukemia (myeloblastic, promyelocytic, myelomonocytic, monocytic and erythroleukemia); chronic leukemia (chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia); and polycythemia vera, lymphoma (Hodgkin's disease and non-Hodgkin's disease), multiple myeloma, Waldenström's macroglobulinemia, and heavy chain disease.

The SNP detection methods can further be useful in determining genetic differences and diagnosis of patients with autoimmune diseases, including but not limited to, insulin dependent diabetes mellitus, multiple sclerosis, systemic lupus erythematosus, Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis, mixed connective tissue disease, primary biliary cirrhosis, pernicious anemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo, gluten-sensitive enteropathy, Graves' disease, myasthenia gravis, autoimmune neutropenia, idiopathic thrombocytopenia purpura, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune infertility, Goodpasture's disease, bullous pemphigoid, discoid lupus, ulcerative colitis, and dense deposit disease.

Genotyping screening methods may also be used screening for specific SNPs not associated with disease. Non-limiting examples include SNPs present in non-coding genomic sequences, or SNPs associated with different human blood groups.

In a preferred aspect of the invention, the methods of the invention may have particular utility in the identification of human SNPs. It is appreciated, however, that the methods described herein will be useful in diagnosing diseases of other mammals, for example, farm animals including cattle, horses, sheep, goat, and pigs, household pets including cats and dogs, and plants including agriculturally important plants and garden plants.

6. EXAMPLE: ANALYSIS OF PCR-FREE GENOTYPE MAPPING

To demonstrate how the methods of the invention are analyzed for the presence of polymorphisms in a DNA sample, the following example is provided. The steps described herein are shown schematically in Figure 2. Methods for preparation and labeling of probes and for carrying out the enzymatic reactions are described in detail in Section 5 herein above. The single nucleotide polymorphism (SNP) of interest in this example is a G/C base-pair polymorphism located at a site where the normal, wild-type (WT) base pair is A/T.

As shown in Figure 2, a peptide-labeled oligonucleotide "positive probe" is synthesized having a G at the SNP site. The probe has two fluorophores, F_1 and F_2 , positioned on the second nucleotide from the 5' end and on the 3' end, respectively, as shown in Figure 2. A second peptide-labeled oligonucleotide is prepared having the same sequence but with a T at the SNP site. This probe is the "negative probe". The negative probe is labeled with the same fluorophores as the positive probe, at precisely the same positions.

In the first nucleic acid sample to be analyzed, the SNP occurs on both copies of the gene. The probes are annealed to aliquots of the target DNA sample in separate reactions. After binding to antibody chips, or a single antibody chip with addressable loci, and samples are digested with mung bean nuclease to remove the overhanging ends. A first reading is taken to detect signals from the two fluorophores. The readout is expressed as a ratio of signal from sample, normalized to a known SNP control sample. Since both the positive and the negative probes retain both fluorophores, the first reading (R_1) for both samples is expected to be 1.0 (see Table 2 below).

Table 2

Positive SNP Probe (G)

Negative WT Probe (T)

	R ₁	R ₂
F1	1.0	1.0
F2	1.0	1.0

	R ₁	R ₂
F1	1.0	1.0
F2	1.0	0

The chip is then immersed in a solution containing S1 nuclease (S1) and Endonuclease V (Endo V), which results in strand scission at mismatched base pairs. The enzymes will not cleave the SNP site of the positive SNP probe, since the target DNA will have the same sequence as the probe. However, the negative probe will be cleaved at the SNP site since the wild type base T of the probe will not pair with the G of the target DNA. Thus, the positive probe retains both fluorophores, whereas the negative probe retains only the F1 fluorophore. A second reading is taken (R₂). Again, as with the first reading, a control sample that has not been treated with nuclease provides a baseline reading. The readout is expressed as the signal relative to a baseline reading from the control sample where the antibody chip was not treated with enzymes, and the expected results are shown in Table 2 below.

To determine whether a SNP is present, a ratio of F2 to F1 is calculated for each probe and each reading, as shown in the top columns of the two charts in Table 3. For the SNP probe the expected F2/F1 ratio is 1.0 for R₁ and 1.0 for R₂. The two ratios are compared. If the first ratio is at least 20% greater than the second ratio then a mismatch is identified. More preferably, the first ratio is at least 25%, or most preferably 35% greater, than the second ratio to identify a mismatch. Here, no mismatch is identified. Thus, the DNA is the same as the probe, *i.e.*, it contains the polymorphic SNP site. To confirm this result, the same calculation is made with the negative WT probe. For the WT probe the expected F2/F1 ratio is 1 for R₁ and 0 for R₂. The two ratios are compared. If the R₁ ratio is at least 20% greater than the R₂ ratio then a mismatch is identified. More preferably, the R₁ ratio is at least 25% greater, or most preferably 35% greater, than the the R₂ ratio to identify a mismatch. Since 1 is greater than 0, a mismatch is identified, the target DNA is not wild type, confirming the result with the positive probe.

Using a target DNA sample with a wild type base at the SNP site, the opposite result is expected to be obtained with both probes. Typical expected results are shown in the lower panels of the charts in Table 3.

Table 3

Positive SNP Probe (G)

Target DNA	R ₁	R ₂
T/A (SNP)	1.0	1.0
G/C (wild type)	1.0	0

Negative WT Probe (G)

Target DNA	R ₁	R ₂
T/A (SNP)	1.0	0
G/C (wild type)	1.0	1.0

The analysis is the same even in experiments using more complex target DNA samples. For example, a DNA sample with a heterozygote genotype, *i.e.*, where the target DNA sample contains one allele with a SNP and the other allele with a WT base, can be mapped using this method. In this case, half the hybrids will have mismatched base pairs at the SNP site in experiments using either the positive SNP probe or the negative WT probe, and thus only half the probes in any sample will be cleaved with S1 and Endo V. Again, the ratio of F2 to F1 is calculated for each probe and each reading, as shown in the middle columns of the two charts in Table 4, below. For the SNP probe the expected F2/F1 ratio is 1.0 for R1 and 0.5 for R2. The two ratios are compared. If the first ratio is at least 20% greater than the second ratio then a mismatch is identified. More preferably, the first ratio is at least 25%, or most preferably 35% greater, than the second ratio to identify a mismatch. Here, a mismatch is identified. The same calculation is made with the negative WT probe. For the WT probe the expected F2/F1 ratio is again 1.0 for R1 and 0.5 for R2. Again, where the R1 ratio is at least 20%, or preferably, 25%, or more preferably, 35% greater than the R2 ratio and a mismatch is identified. The target DNA is therefore neither entirely wild type nor SNP in genotype. Because the F2/F1 ratio for both probes is again 1.0 for R1 and 0.5 for R2, a heterozygosity is identified. The result can be confirmed using control target DNA samples with homozygous WT and SNP alleles. The expected results for all three possible genotypes are shown in Table 4 below.

Table 4

Positive SNP Probe (G)

Allele ₁	Allele ₂	R ₁	R ₂
G/C (wild type)	G/C (wild type)	1.0	1.0
G/C (wild type)	T/A (SNP)	1.0	0.5
T/A (SNP)	T/A (SNP)	1.0	0

Negative WT Probe (A)

Allele ₁	Allele ₂	R ₁	R ₂
G/C (wild type)	G/C (wild type)	1.0	0
G/C (wild type)	T/A (SNP)	1.0	0.5
T/A (SNP)	T/A (SNP)	1.0	1.0

Similar analysis is used to compare more complex populations of target DNA samples. For example, in one embodiment of the invention, the multiplex PCR-free SNP detection method is used to associate a particular disorder with a genotype. In this case, two target DNA samples are compared – a sample comprising a plurality of DNA samples from a population afflicted with the disorder of interest, and a control sample obtained from a non-afflicted population. The analysis, as described above, is used to detect differences in a plurality of SNP sites between the two populations. The presence of a SNP site in the afflicted population, but absent in the control population allows correlation of said SNP site with the disorder.

The invention described and claimed herein is not to be limited in scope by the specific embodiments herein disclosed since these embodiments are intended as illustration of several aspects of the invention. Any equivalent embodiments are intended to be within the scope of this invention. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims. Throughout this application various references are cited, the contents of each of which is hereby incorporated by reference into the present application in its entirety.